

Name \_\_\_\_\_

A2CC Intro to Stat – homework

1. Scientists randomly select ten groups from a population of men over 50 years old. They calculate the mean weights of each of these groups. The variability between these means can be best attributed to
  - (1) measurement variability
  - (2) natural variability
  - (3) induced variability
  - (4) sampling variability
2. Max and Daniel are measuring the amount of time it takes for a ball to roll down a ramp at different heights. For each trial, both Max and Daniel take turns rolling the ball and working the stop watch. They do this in order to quantify which of the following sources of variability?
  - (1) measurement variability
  - (2) natural variability
  - (3) induced variability
  - (4) sampling variability
3. Which of the following scenarios would be an attempt to quantify induced variability?
  - (1) a phone survey of political preferences during election season
  - (2) multiple random samples of products from an assembly line to check for defects
  - (3) random assignment of people to a control group and a group taking a drug to lower cholesterol
  - (4) recording the variability in the measurement of a soil sample's weight by the same machine
4. Which of the following research questions would involve collecting data through a survey?
  - (1) Watching people exit a grocery store to see the percent who use reusable bags.
  - (2) Assigning people to two groups to see the effect of a particular amount of sleep.
  - (3) Calling people on the telephone to see if they will be voting in the upcoming election.
  - (4) Dropping salt cubes into two different liquids to determine which dissolves faster.
5. In which of the following cases would an observational study be necessary as compared to an experimental study?
  - (1) The study of how increased nutrient levels affect plant growth.
  - (2) The study of how educational levels affect median household income.
  - (3) The study of how a vaccine affects the percent of mice that get a particular disease.
  - (4) The study of how noise level affects the sleep patterns of volunteers in a sleep study.
6. Which of the following formulas, written in summation notation, would represent the mean of the data set  $\{x_1, x_2, \dots, x_n\}$ ? Explain your choice.

(1)  $\sum_{i=1}^n x_i$

(2)  $n \sum_{i=1}^n x_i$

(3)  $\frac{1}{n} \sum_{i=1}^n x_i^2$

(4)  $\frac{1}{n} \sum_{i=1}^n x_i$
7. The standard deviation of a population characteristics measures
  - (1) The difference between the maximum and minimum values.
  - (2) The difference between the third quartile and first quartile values.
  - (3) The average distance a data value is away from the mean.
  - (4) The average distance a data value is away from the median.
8. The interquartile range of the data set  $\{4, 7, 10, 13, 18, 22, 30\}$  is
  - (1) 15
  - (2) 7
  - (3) 18
  - (4) 10

9. In an experimental study, a lab wanted to divide volunteers into two groups to determine the effect of a particular phone app to help make people more punctual (on time). The 50 volunteers in the study will be assigned to either a group of 25 who use the app for a week or a group of 25 who do not use the app. The participants were asked to come to a lab to receive the app (or not) at 10:00 am on a Monday. Answer the following questions:

(a) Why would those performing the study *not* want to assign the participants in the two treatments (groups) based on who showed up to the study session first?

(b) Propose a way to use a random number table to generate a simple random selection that eliminates the bias that you discussed in part (a).

10. If you were trying to conduct a survey of political preferences for likely voters in an upcoming election and decided to dial 1,000 randomly generated land-line phone numbers (not cell), why might this still introduce **bias** into the sampling?

11. The heights of the 15 players on the Arlington boys' varsity basketball team are given below in inches.

66, 67, 68, 68, 70, 72, 72, 73, 74, 75, 75, 75, 76, 77, 79

(a) Find the mean and standard deviation of this data set. Use the population standard deviation. Round both to the nearest *tenth*.

(b) Determine the proportion of the population that falls within one standard deviation and within two standard deviations of the mean. State your values in decimal form.

One standard deviation from the mean:

(c) Use the random number table for this lesson to pick a random sample of five players from this list. Do this by picking a random two digit column along the page. Scan down the column until you have picked 5 random integers that fall from 1 to 15. Write down your sample and calculate its mean.

Two standard deviations from the mean: